

CSA1618 DWDM-DE

EXPERIMENT-22

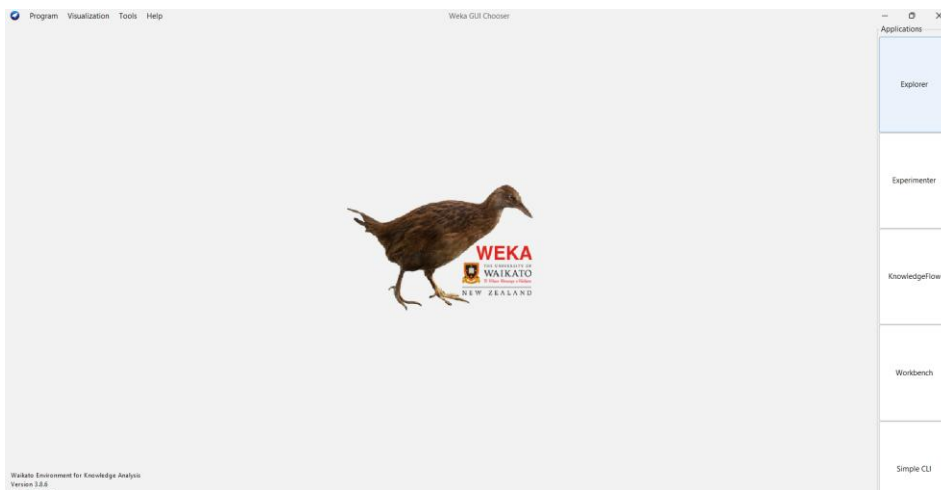
DATA SEGMENTATION BY K-MEANS CLUSTER USING WEKA AND R-TOOL

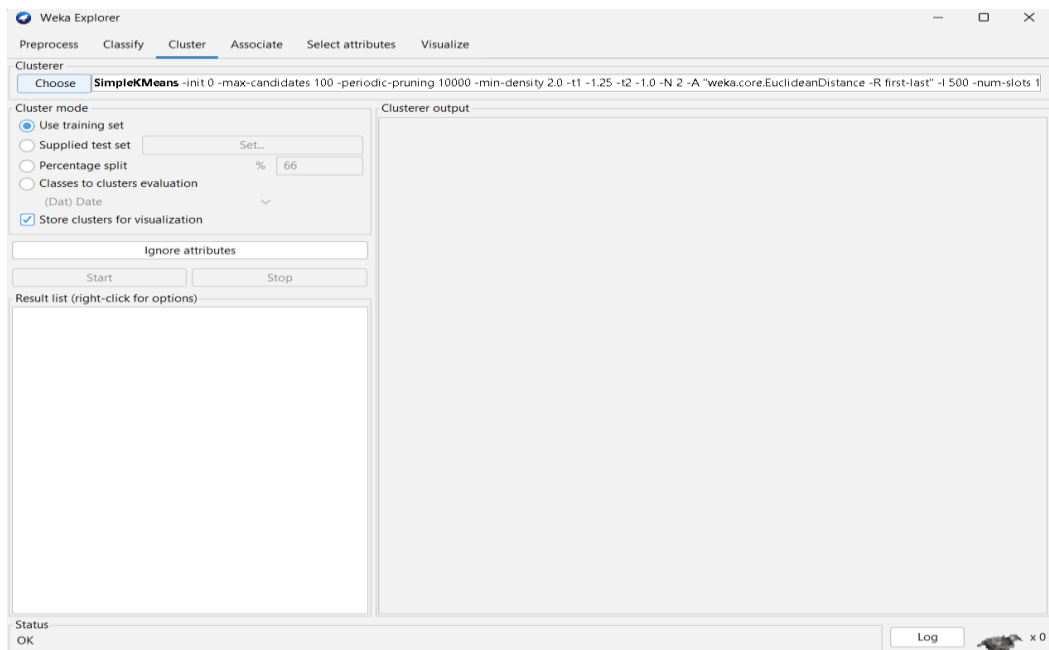
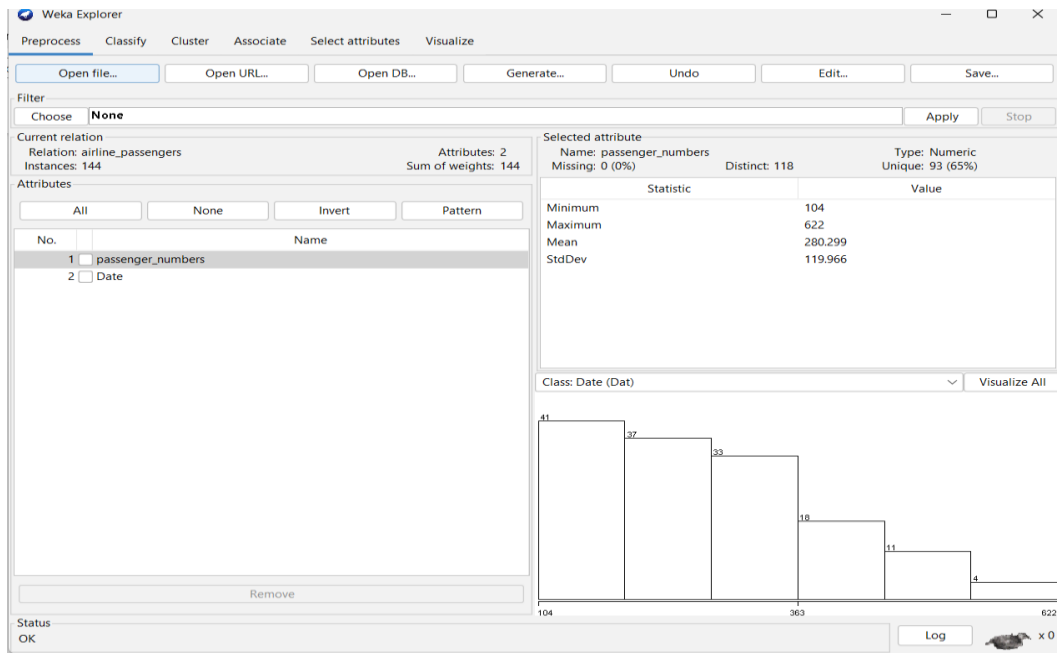
AIM:

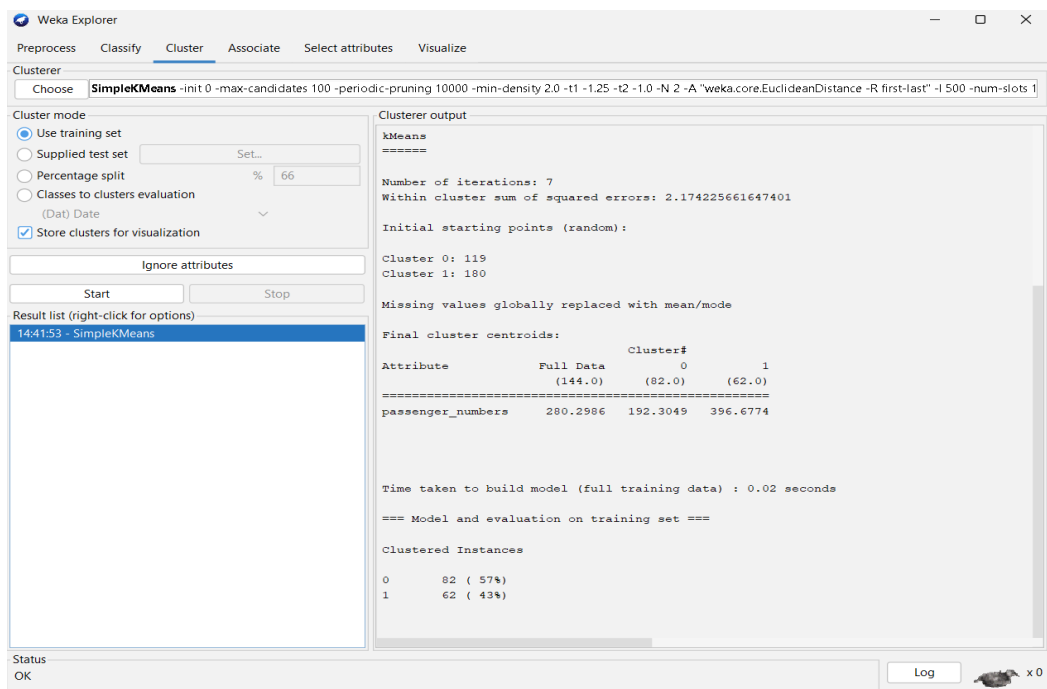
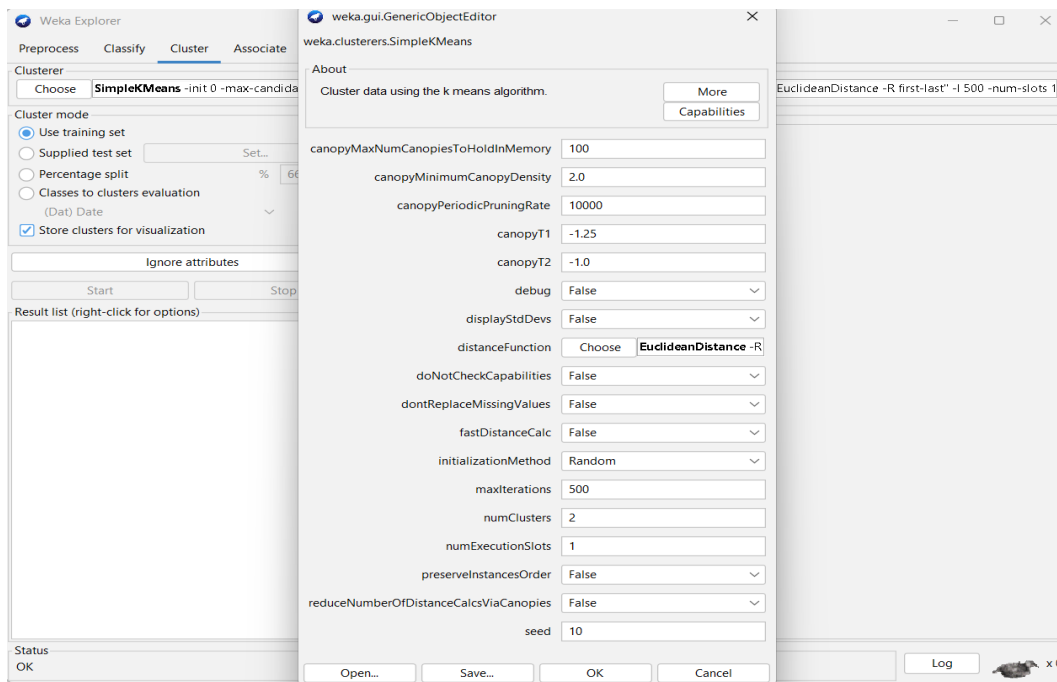
To create DataSegmentation by k-means cluster using weka and R-tool.

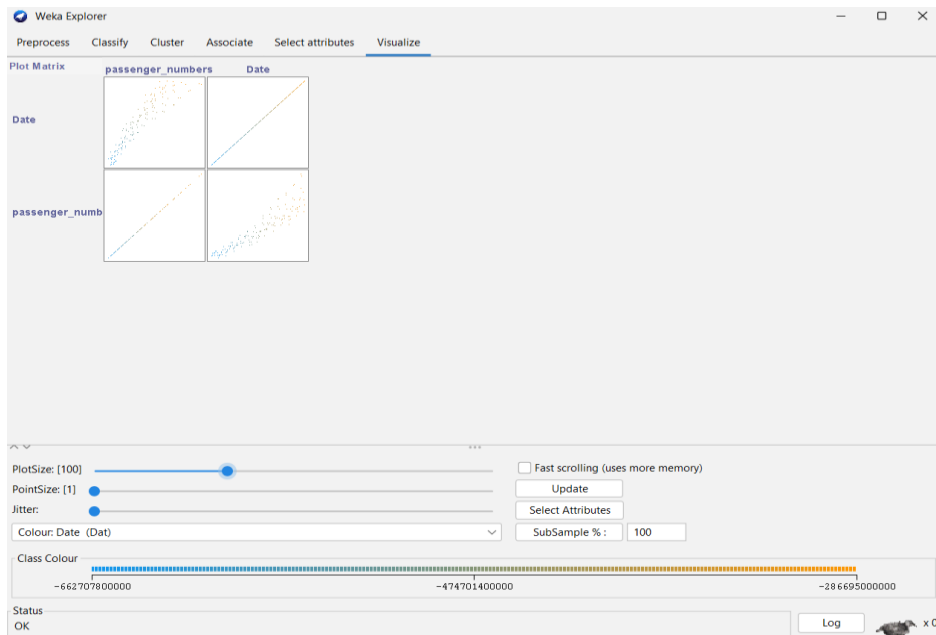
PROCEDURE:

1. Download and install WEKA.
2. Open WEKA and Choose "Explorer" from the main menu.
3. Under Preprocess, Click on the open file button and select the dataset.
4. Click on the "Cluster" tab. In the Cluster mode section, select "Use training set".
5. Click "Choose" (next to the cluster algorithm) and Select SimpleKMeans (under weka.clusterers).
6. Click on "SimpleKMeans" to configure it - Set "numClusters" to the desired number of clusters (e.g., 2, 3, 5). Set "Seed" (random initialization, e.g., 10). Choose "Euclidean distance" (default).
7. Click "OK" and then "Start" to run clustering.
8. WEKA will display cluster assignments and statistics. Click "Visualize" to see how the clusters are distributed. Save the file.









OBSERVATION:

K-means

=====

Number of iterations: 7

Within cluster sum of squared errors: 2.174225661647401

Initial starting points (random):

Cluster 0: 119

Cluster 1: 180

Missing values globally replaced with mean/mode

Final cluster centroids:

	Cluster#	
Attribute	Full Data	0 1
	(144.0)	(82.0) (62.0)

=====

passenger_numbers 280.2986 192.3049 396.6774

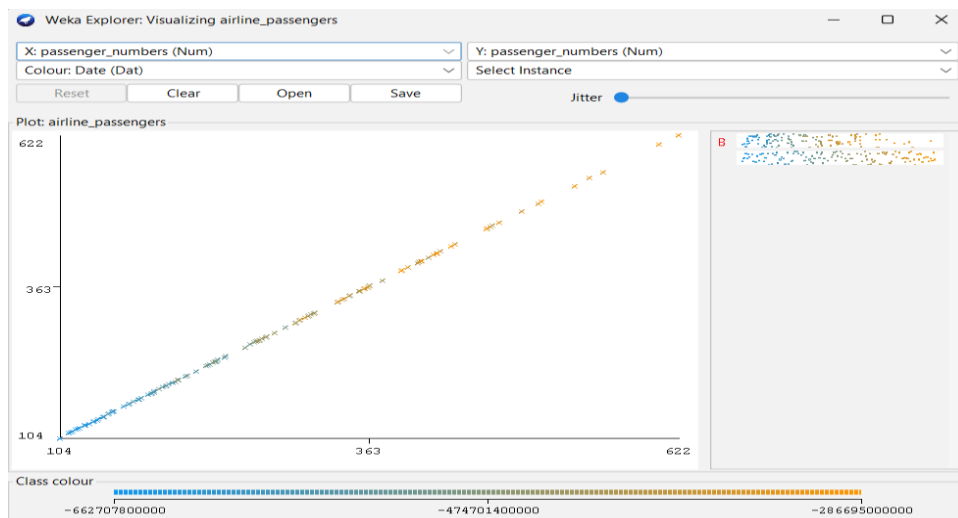
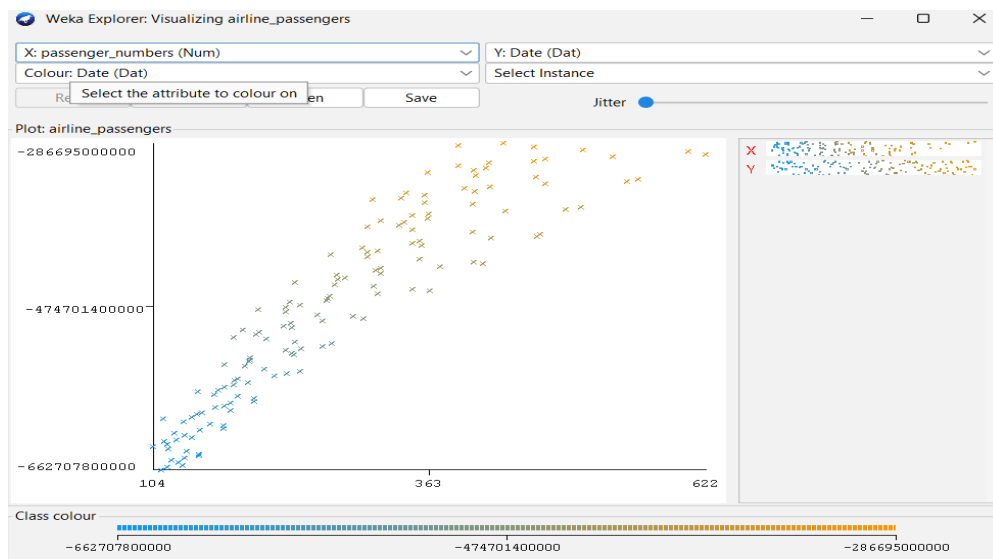
Time taken to build model (full training data) : 0.02 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 82 (57%)
1 62 (43%)

PLOT:



RESULT:

Thus, the K-means clustering analyzing using the weka tool has been successfully completed. In case of weka tool, the change in seed values lead to the decrease in the number of iterations.